11) Japanese Patent Application

Laid-open (KOKAI) No. 6-236197

(43) Laid-opened Date: August 23, 1994

(54) Title of the invention:

PITCH PATTERN GENERATOR

(21) Application Number: 5-180056

(22) Filing Date: June 24, 1993

(71) Applicant: RICOH CO LTD

(72) Inventor: Tetsuya SAKAYORI

---

(54)[Title of the Invention] Pitch pattern generator

(57)[Abstract]

[Objects]

To be suitable for application to an apparatus as in rule-based synthesis that synthesizes any voice and to be able to generate pitch patterns that are more natural than conventional ones.

[Constitution]

A pattern storing unit 1 stores therein unit patterns cut out on an appropriate pattern basis from pitch patterns of human speeches; a pattern searching unit 2 searches, based on a predetermined input attributes, unit patterns conforming to the input attributes from the pattern storing unit 1; a pattern disposing unit 3 disposes on a chronological axis unit patterns obtained by searching; and a pattern transforming unit 4 transforms unit patterns disposed

on the chronological axis to generate pitch patterns. Since the pitch patterns that a man has spoken are directly used, such pitch patterns that will be able to express natural variations, fluctuation, nuance and the like, and are difficult for rule-based generation, can be generated.

[Claims]

[Claim 1] A pitch pattern generator characterized by comprising:

pattern storing means for storing unit patterns cut out in an appropriate unit basis from pitch patterns of human speeches;

pattern searching means for searching, based on a predetermined input attribute or input linguistic information, unit patterns which conform to the input attribute or input linguistic information;

pattern disposing means for disposing the unit patterns obtained by searching onto a chronological axis; and

pattern transforming means for transforming the unit patterns disposed on the chronological axis.

[Claim 2]

The pitch pattern generator according to claim 1, characterized in that as said unit pattern, at least one of a pattern, containing one ascending portion and one descending portion respectively of an accent phrase, cut out as a unit; a pattern, containing one ascending

portion or one descending portion of an accent phrase, cut out as a unit; a pattern, with a pitch pattern section containing a flat portion of an accent phrase, cut out as a unit; a pattern, with a portion showing characteristic pitch shift at the end of a phrase, cut out as a unit; a pattern, with a pitch pattern section showing a particular range of shift rates, cut out as a unit; and a pattern, with a pitch pattern section showing a particular range of curvature rates, cut out as a unit, is used.

[Claim 3]

The pitch pattern generator according to claim 1, characterized in that said pattern storing means stores the unit patterns cut out or the unit patterns which are cut-out and vector quantized.

[Claim 4]

The pitch pattern generator according to claim 1, characterized in that said pattern searching means is arranged to use as an input attribute, at least one of information among phonological environmental information over a plurality of morae in the vicinity of a pitch shifting portion, information related to prominence within a unit pattern or in the vicinity thereof and information of intonation style of a phrase where a unit pattern is contained.

[Claim 5]

The pitch pattern generator according to claim 1,

characterized in that said pattern searching means is arranged to determine an optimum unit pattern with neural networks based on the input attributes.

[Claim 6]

The pitch pattern generator according to claim 1, characterized in that said pattern disposing means is arranged to determine, a location of a unit pattern on the chronological axis, using as input information at least one among the location of accent phrase boundary on the chronological axis, the location of accent core on the chronological axis, the location of boundary such as phonology, syllable and mora at the end of a phrase on the chronological axis and the phonological environment information in the vicinity of pitch shifting portion.

[Claim 7]

The pitch pattern generator according to claim 1, characterized in that said pattern transforming means transforms a flat portion to expand and contract in the direction of chronological axis in the case where the unit pattern has flat portion boundary information in an accent phrase, and transform an ascending portion and a descending portion to expand and contract in the direction of pitch axis in the case where the unit pattern has boundary information of an ascending portion and a descending portion in an accent phrase.

[Claim 8]

The pitch pattern generator according to claim 1, characterized by further comprising pattern connecting means for connecting the unit patterns subject to shifting with said pattern transforming means to generate a series of phrase patterns.

[Claim 9]

The pitch pattern generator according to claim 8, characterized in that said pattern connecting means is arranged to bring the unit patterns into connection with interpolation by line, or to bring the unit patterns into connection with interpolation by three dimensional curve.

[Claim 10]

The pitch pattern generator according to claim 1, characterized in that, as the unit patterns stored in said pattern storing means, pitch patterns characteristically appearing in a particular linguistic situation are used.

[Claim 11]

The pitch pattern generator according to claim 10, characterized in that, as one of the unit patterns, pitch patterns characteristically appearing in a particular location in a series of phrases such as breath group and sentences are used.

[Claim 12]

The pitch pattern generator according to claim 10, characterized in that as one of the unit patterns,

pitch patterns characteristically appearing in a particular accent type such as an initial rise type and a plateau type are used.

[Claim 13]

The pitch pattern generator according to claim 10, characterized in that as one of the unit patterns, pitch patterns which characteristically appear in a particular accent levels such as in a portion provided with prominence and in a portion where a subsidiary accent core appears are used.

[Claim 14]

The pitch pattern generator according to claim 10, characterized in that as one of the unit patterns, words which are uttered with their own characteristic pitch patterns such as in an interjection, an end particle, an onomatopoeic word, and imitative word and a shout are used, or phrasal pitch patterns are directly used.

[Claim 15]

The pitch pattern generator according to claim 10, characterized in that as one of the unit patterns, characteristic pitch patterns for periodically appearing in a line of numbers and a verse are used.

[Claim 16]

The pitch pattern generator according to claim 10, characterized in that as one of the unit patterns, characteristic pitch patterns expressing particular

intentions mainly by way of meter independent from words and phonology are used.

[Claim 17]

The pitch pattern generator according to claim 1, characterized in that said pattern storing means stores the unit patterns subject to approximation with lines.

[Claim 18]

The pitch pattern generator according to claim 17, characterized in that the unit patterns subject to the approximation with lines are expressed with expression data with the pitch at the starting point and the inclination between respective section of each line and the line, and stored in said pattern storing means.

[Claim 19]

The pitch pattern generator according to claim 18, characterized in that the inclination of each line is expressed by step values as time required for the pitch cycle to change by a unit quantity, and in a case where said pattern searching means searches expression data of the unit patterns stored in said pattern storing means and reproduces the unit patterns based on the expression data, said pattern searching means implements process to increase the pitch cycle by unit period in case of the step values being positive values and decrease the pitch cycle by unit period in case of the step values being negative values.

[Claim 20]

The pitch pattern generator according to claim 19, characterized in that said pattern transforming means adds a constant bias to the step value representing inclination of each line so as to provide the unit patterns with smooth shifting in the pitch direction.

[Claim 21]

The pitch pattern generator according to claim 20, characterized in that said pattern transforming means treats a pitch renewal cycle with complement expression of "2" and, in a case where the absolute value of the step value as a result of application of bias transcends the maximum value, regard this as the maximum absolute value with an opposite sign.

[Claim 22]

The pitch pattern generator according to claim 20 or 21, characterized in that in a case where corresponding relationship between the pitch shifting quantity and the bias is obtained in advance, said pattern transforming means implements shifting in the pitch direction of the unit pattern with the corresponding relationship.

[Claim 23]

The pitch pattern generator according to claim 1, characterized by further comprising rule storing means for storing rules including conditions corresponding to the unit patterns stored in said pattern storing means, wherein said pattern searching means is arranged to

look into rule conditions stored in said rule storing means in each input linguistic information unit such as morae etc. and to search the unit patterns fulfilling the rule conditions from said pattern storing means.

[Claim 24]

The pitch pattern generator according to claim 23, characterized in that the rule conditions include ranges of values of a plurality of variables representing characteristics of linguistic information unit, and said pattern searching means is arranged to adopt the rule, in a case where for all the variables included in the rule conditions, the values of the input linguistic information unit are within that range, and thereby to search the unit patterns.

[Claim 25]

The pitch pattern generator according to claim 24, characterized in that as the variables included in the rule conditions, locations from the phrase top or the phrase end of the linguistic information unit are used, classification of boundary with the phrase adjacent to the phrase containing the linguistic information unit is used, accent information of the phrase containing the linguistic information unit is used or phonological information of the linguistic information unit is used.

[Claim 26]

The pitch pattern generator according to claim 23, characterized in that a plurality of unit patterns are

brought into correspondence to one rule, and, in the case where a plurality of unit patterns fulfilling the rule condition are searched, fluctuation is provided so that one unit pattern is selected.

[Detailed Description of the Invention]

[0001]

[Field of the Invention]

The present invention relates to a pitch pattern generator which is used for an apparatus for speech-synthesis by rule and generates chronological patterns of pitch frequency, that is, pitch patterns.

[0002]

[Prior Art]

Conventionally, synthesized sound from an apparatus for speech-synthesis by rule is mechanical and monotonous, and is obviously different from human voice, lacking naturalness to give rise to incongruity to a listener. In order to obtain synthesized sounds similar to natural voice, apparatuses introduced, for example, in Japanese Patent Laid-Open No. 2-197897 (hereinafter referred to the conventional example 1) or Japanese Patent Laid-Open No. 3-139699 (hereinafter referred to the conventional example 2), have been proposed.

[0003]

That is, a pitch pattern generator, which is proposed in a prior art example 1, stores an ascending

pattern of pitch frequency on each kind of phoneme in a mora where the pitch frequency ascends in accordance with an accent type of a word or a word group; stores a descending pattern of pitch frequency on each kind of phoneme in a mora where the pitch frequency descends in accordance with an accent type of a word or a word group; reads out an ascending pattern of the above described stored pitch frequency in accordance with the kind of phoneme in a mora where the pitch frequency ascends in accordance with an accent type of the synthesized word or a word group; reads out a descending pattern of the above described stored pitch frequency in accordance with the kind of phoneme in a mora where the pitch frequency descends in accordance with an accent type of the synthesized word or a word group; and brings the ascending pattern of the read out pitch frequency and the descending pattern of the read out pitch frequency into connection to generate a pitch patter of the whole of a word or a word group, and the apparatus is intended to generate a pitch pattern close to a pitch pattern of a natural voice in a mora with the pitch frequency level which changes.

[0004]

In addition, among voice editing synthesizers that stores a voice of prosodic word as a unit and can control pitch patterns at will, a voice editing synthesizer which is proposed in a prior art example 2,

stores only accent components as to pitches at the time of storing each prosodic word and generates the phrase components corresponding to that paragraph at the time of synthesizing the paragraph and overlaps the accent components onto the phrase components and thereby generates a pitch pattern, and the voice editing synthesizer is intended to generate such a pitch pattern as to make the intonations over the entire paragraph natural.

[0005]

[Problems to be Solved by the Invention]

Thus, any apparatus described in the conventional example 1 and the conventional example 2 is intended to obtain synthesized sounds similar to natural voice. However, with the pitch pattern generator of the conventional example 1, the patterns for use are limited to such parts that their pitches ascend or descend according to their accent types, therefore leaving the other parts to remain unnatural. In addition, since the pitch frequency of this pattern is categorized only in accordance with the kind of phonemes in a position changing from "high" to "low" level or in the opposite direction, the same pattern is used in case of changes involving the same phonology, resulting in making a monotonous voice, and moreover without considering surrounding phonological environments or existence of prominence, therefore

leaving unnatural impression, to give rise to a problem.

[0006]

In addition, the device of the conventional example 2, which relates to a voice editing synthesizer, needs to store accent components corresponding to all prosodic words in use, but since prosodic words generally consists of combinations of words and phrases, it is actually impossible to cover all prosodic words to store. In addition, changes in positions within a paragraph, in a context, an intonation, a prominence, or speech speeds, will not allow direct use. Accordingly, an application to such device that synthesizes any voices as in rule-based synthesis is difficult, giving rise to a problem.

[0007]

An object of the present invention is to provide a pitch pattern generator which is suitable to application to a device to solve such a problem, and synthesize any voices as in rule-based synthesis as well, generate more natural pitch patterns than conventional ones.

[0008]

[Means for Solving the Problems and Operation]

In order to achieve the above described objects, the invention according to claim 1 comprises pattern storing means for storing unit patterns cut out on an appropriate unit basis from pitch patterns of human

speeches; pattern searching means for searching, based on a predetermined input attribute, unit patterns which conform to the attribute; pattern disposing means for disposing the unit patterns obtained by searching onto the chronological axis; and pattern transforming means for transforming the unit patterns disposed on the chronological axis. Thus, the pitch patterns that a human being has spoken can be directly used to generate such pitch patterns that will be able to express natural variations, fluctuation, nuance and the like that are difficult for rule-based generation.

[0009]

In addition, in the invention according to claim 2, in case of cutting out the one containing one ascending portion and one descending portion of accent phrase respectively as a unit that is to be used as a unit pattern, pitch patterns with the accent phrase patterns in nearly full preservation being important units of pitch patterns can be generated. In addition, in case of using the one containing one ascending portion or one descending portion of accent phrase as a unit pattern, pitch patterns can be generated with a simple treatment. In addition, in case of using unit patterns containing flat portions of accent phrase, fluctuation and the like in the constant domain can be preserved and natural pitch patterns can be generated. In addition, in case of using characteristic pitch

variations in the end of phrases as unit patterns, various intonations which it is difficult to generate with rules can be expressed. In addition, in case of using patterns showing percentage of change of pitch within a particular range as unit patterns, rough form in shifting portions and fluctuation etc. in the constant domain portions can be preserved so that various and natural pitch patterns which it is difficult to generate with rules can be generated. In addition, in case of using patterns showing pitch curvatures within a particular range as unit patterns, the other portions can be interpolated with straight lines, process quantity can be reduced.

[0010]

In addition, in the invention according to claim 3, the unit patterns subject to vector quantization are stored, and when this is used, the volume of pattern storing means can be dramatically reduced so that cost reduction as well as miniaturization in the device and increase in processing speed can be achieved.

[0011]

In addition, in the invention according to claim 4, in case of searching unit patterns based on phonology environments over a plurality of morae in the vicinity of pitch shifting portions, the reproducing nature of local shifts of patterns due to phonological environments is improved so that more precise pitch

patterns can be generated. In addition, in case of searching unit patterns based on prominence information, intonations in synthesized voices which conventionally tend to become monotonous are made well-modulated so that intelligibleness and naturalness can be improved. In addition, in case of searching unit patterns based on intonation styles, intonations which it is difficult to generate with rules can be accurately reproduced.

[0012]

In addition, the invention according to claim 5 is to search unit patterns with neural networks, thereby enabling reproduction of patterns close to speech patterns learned from a human being so that naturalness can be improved.

[0013]

In addition, in the invention according to claim 6, in case of disposing unit patterns with locations of accent phrase boundary with pattern disposing means, the ascend starting points of accent phrases can be reproduced to an approximately full extent so that intelligibleness can be improved. In addition, in case of disposing unit patterns with locations of accent cores, the descend starting points of accent phrases can be reproduced to an approximately full extent so that intelligibleness can be improved. In addition, in case of disposing unit patterns with boundary location in the end of phrases such as phonology, syllables and

morae, etc., intonations can be expressed efficiently so that naturalness can be improved. In addition, by disposing the unit patterns with phonological environment information in the vicinity of pitch shifting portions, chronological structure of the pitch patterns can be reproduced more precisely so that naturalness can be improved.

[0014]

In addition, in the invention according to claim 7, pattern shifting means is to expand and contract in the direction of chronological axis in the plateau portions in the accent phrases so that, as for the pitch shifting portions, without any shift in the direction of chronological axis, the chronological structure of pitch shifting portions can be preserved and stable natural pitch patterns can be generated. In addition, the ascending portions and the descending portions in the accent phrases are to shift in the direction of pitch axis so that, as for the plateau portions, without any shift in the direction of the pitch axis, height or fluctuation etc. of pitch constant domain portion can be preserved, pattern expressions including prominence and the like become feasible and naturalness can be improved.

[0015]

In addition, in the invention according to claims 8 and 9, pattern connecting means for connecting the

unit patterns subject to shifting with pattern transforming means to generate a series of phrase patterns are further provided so that, in case of interpolating with straight lines interconnection between unit patterns in the pattern connecting means, process quantities can be reduced and increase in processing speed as well as cost reduction can be achieved while, in case of interpolating with three dimensional curve interconnection between unit patterns, rough shape similar to human pitch shifts can be expressed and natural pitch patterns can be generated.

[0016]

In addition, the inventions according to claims 10 to 16 are characterized by using pitch patterns characteristically appearing in a particular linguistic situation as unit patterns stored in the pattern storing means. Thereby, from linguistic information being input information of speech-synthesis by rule, patterns can be searched heuristically so that efficient and effective pitch pattern generation can be realized.

[0017]

In addition, the inventions according to claims 17 and 18 are to arrange the pattern storing means to store the unit patterns subject to approximation with line and, at this time, the unit patterns subject to approximation with line are expressed with expression

data with the pitch at the starting point and the inclination between respective section of each line and the line, and stored in the pattern storing means, thereby data quantities of the unit patterns, memory volume and process volume can be reduced and cost reduction as well as increase in processing speed can be achieved. Moreover, shifting in the pitch direction of the unit patterns is only increase or reduction of starting pitch, and expansion and contraction in the chronological direction can be simply realized respectively with increase or reduction of interval length, thereby the shift of the unit pattern can be realized with simple process so that increase in processing speed as well as functional improvement can be achieved.

[0018]

In addition, in the invention according to claim 19, inclination of each of the above described line is expressed by step values as time required for the pitch cycle to change by a unit quantity, and in this case, when the pattern searching means searches expression data of the unit patterns stored in the pattern storing means and reproduces the unit patterns based on the expression data, the pattern searching means implements process to increase the pitch cycle by unit period in case of the step values being positive values and decreases the pitch cycle by unit period in case of the

step values being negative values. Thereby, extremely simple configurations with counters and comparators etc. can reproduce the unit patterns approximated with lines so that actual pitch patterns become obtainable.

[0019]

In addition, in the invention according to claim 20, the pattern transforming means adds a constant bias to the step value representing inclination of each line so as to provide the unit patterns with smooth shifting in the pitch direction. Thereby, the pattern transforming process that will become necessary for connection between unit patterns or provision of prominence can be realized with extremely simple configurations only with an adding machine.

[0020]

In addition, in the invention according to claim 21, the pattern transforming means treats the step values with complement expression of "2" and, in the case where the absolute value of the step value as a result of application of bias has transcended the maximum value, regards this as the maximum absolute value with opposite sign. Thereby, a process to implement pattern transforming continuously and without compulsiveness from right ascending pattern to right descending pattern or vice versa can be realized simply.

[0021]

In addition, in the invention according to claim

22, in the case where corresponding relationship between the pitch transforming quantity and the bias is obtained in advance, the pattern transforming means implements shifting in the pitch direction of the unit pattern with the corresponding relationship. Thus, the bias that will become necessary in connection between unit patterns as well as prominence application and the like can be obtained extremely easily without necessity of multiplication and division with a large amount of process quantity so that increase in processing speed can be achieved.

[0022]

In addition, in the invention according to claim 23, moreover, the rule storing means for storing rules including conditions corresponding with the unit patterns stored in the pattern storing means is provided so that the pattern searching means is arranged to look into the rule conditions stored in the rule storing means in each input linguistic information unit such as morae etc. and thereby search the unit patterns fulfilling the rule conditions from the pattern storing means. Adopting rule searching as the unit pattern searching process, maintenance such as addition, deletion and change etc. in unit patterns will become easy.

[0023]

In addition, in the invention according to claim

24, the rule conditions include ranges of values of a plurality of variables representing characteristics of linguistic information unit, and the pattern searching means is arranged to adopt the rule when for all the variables included in the rule conditions the values of the input linguistic information unit are within that range, and thereby to search the unit patterns. Making the rule conditions include a range of values of a plurality of variables representing characteristics of linguistic information unit, the rule searching can be implemented with routine process. In addition, it will become easy to separate the rules as external data so that maintainability of the rules themselves can be improved.

[0024]

In addition, in the invention according to claim 25, as the variables included in the rule conditions, locations from the phrase top or the phrase end of the linguistic information unit are used, classification of boundary with the phrase adjacent to the phrase containing the linguistic information unit is used, accent information of the phrase containing the linguistic information unit is used or phonological information of the linguistic information unit is used, and usage of locations from the phrase top or the phrase end of the linguistic information unit as the variables included in the rule conditions will enable

expression of nature of pitch patterns continuously changing in accordance with locations in the phrase such as gradual descend from the phrase top to the phrase end. In addition, usage of classification of boundary with the phrase adjacent to the phrase containing the linguistic information unit will enable expression of pitch pattern characteristics in the phrase boundary locations such as direction on ending impression of a statement by dropping the statement end in particular lower than the period position, and dealing with various intonation styles such as questioning statement. In addition, usage of accent information of the phrase containing the linguistic information unit will enable expression of pitch pattern characteristics such as characteristics according to accent type such as the initial rise type constituting a mountain leftward-inclined than the middle rise type, etc. and characteristics according to prominence or subsidiary accent and the like. In addition, usage of phonological information of the linguistic information unit will enable expression of pitch pattern characteristics locally observed in a particular phonology or influence to phonology sustaining chronological length and the like.

[0025]

In addition, in the invention according to claim 26, a plurality of unit patterns are brought into

correspondence with one rule, and, in the case where a plurality of unit patterns fulfilling the rule condition are searched, fluctuation is provided so that one unit pattern will be selected. Thereby, uniform and mechanical pitch patterns will become avoidable and more natural pitch patterns can be obtained.

[0026]

[Embodiments]

Embodiments of the present invention will be described with reference to drawings as follows. Figure 1 is a block diagram of an example of a pitch pattern generator related to the present invention. The pitch pattern generator of the present embodiment comprises a pattern storing unit 1 for storing unit patterns cut out from pitch patterns of natural voices in a predetermined unit, a pattern searching unit 2 for searching, based on the predetermined input attribute, unit patterns in conformity to the input attribute, a pattern disposing unit 3 for disposing onto the chronological axis unit patterns obtained by searching, a pattern transforming unit 4 for transforming unit patterns disposed on the chronological axis and a pattern connecting unit 5 for connecting unit patterns subject to shifting to generate patterns of a series of phrases.

[0027]

Here, the unit patterns stored in the pattern

storing unit 1 undergo cutting out from pitch patterns of natural voices (that is, human speech) in a way as shown in Figure 2, Figure 3, Figure 4 or Figure 5. That is, in Figure 2, a unit pattern UNP is arranged to undergo cutting out as a unit DM_0 from the starting point of the ascending portion AU of an accent phrase AK of a pitch pattern PP uttered in a statement to the ending point of the descending portion AD via the plateau portion AF. In addition, in Figure 3, a unit pattern UNP is arranged to undergo cutting out as a unit DM_1 and a unit DM_2, respectively, for the section from the staring point of the ascending portion AU of the accent phase AK of the pitch pattern PP uttered to the ending point of the subsequent plateau portion AF and the section from the staring point of the descending portion AD to the ending point of the subsequent plateau portion AF. In addition, in Figure 4, a unit pattern UNP is arranged to undergo cutting out to cover only the plateau portion AF as a unit DM_3. In addition, in Figure 5, a unit pattern UNP is arranged to undergo cutting out to cover a characteristic pitch shifting portion expressing an intonation of the ending of a sentence as a unit DM_4. Here, in Figure2 to Figure 5, the plateau portion AF of an accent phrase refers to both of a convex portion from the ending point of the ascending portion AU of a pitch to the starting point of the descending portion

AD and the opposite concave portion. In addition, cut out section as shown in Figure 2 to Figure 5 is implemented in connection with attribute information such as accent core location, but otherwise, in spite of being not shown, a unit pattern may be cut out with a pitch pattern interval representing pitch shifting rate over a particular range as a unit or a unit pattern may be cut out with a pitch pattern interval representing a curvature within a certain range as a unit.

[0028]

Thus, in the pattern storing unit 1, various unit patterns cut out from statement voices in any of the above described ways are stored, or unit patterns vector-quantized with a code book etc. prepared in advance are stored as a database.

[0029]

In addition, the pattern searching unit 2 searches a unit pattern conforming to a predetermined input attribute from the pattern storing unit 1, and as a way for this search, a method of adding to each unit pattern stored in the pattern storing unit 1 attribute information at the time of cutting this out to select a unit pattern having attribute information most similar to the input attribute. Moreover, in a case where, at the time of cutting out a unit pattern, a unit pattern is vector-quantized with a code book prepared in

advance and is stored in the pattern storing unit 1, such a method may be considered that the code of the vector-quantized unit pattern is regarded as an output and a predetermined input attribute is input to a neural network which is made to learn by such a master data that inputs the attribute at that time to search a unit pattern.

[0030]

In addition, the predetermined input attribute used in the pattern searching unit 2 is information considered to influence the pitch pattern in such information that is provided to rule-based synthesizer, and there is an accent type and the like as a representative, but otherwise the following information is added to the input attribute. That is, conventionally, it is known that classification of phonology in the vicinity before or after the pitch shifts gives rise to a local shift of the pitch, thus considering usage of this information in pattern search effective, this is added to the input attribute. In addition, conventionally, it is known that the portion where prominence is added gives rise to shift in height of accent and the like, thus considering usage of this information in pattern search effective, this is added to the input attribute. In addition, in case of selecting patterns as shown in Figure 5, information of phrase intonation style is necessary, thus considering

usage of this information in pattern search effective, this is added to the input attribute.

[0031]

In addition, the pattern disposing unit 3 is arranged to determine a location of the unit pattern on the chronological axis, with the location of accent phrase boundary on the chronological axis as input information, with the location of accent core on the chronological axis as input information, or with the location of boundary such as phonology, syllable and mora at the end of a phrase on the chronological axis as input information, or with the phonological environment information in the vicinity of pitch transforming portion as input information.

[0032]

In addition, the pattern transforming portion 4 corresponds to classification and the like of the used unit pattern so as to make the ending point of each unit pattern correspond to the starting point of the successive unit pattern, so that any one or two or all of the ascending portion AU, the plateau portion AF and the descending portion AD of an accent phrase are arranged to undergo transformation so as to extend or contract in the direction of chronological axis and/or in the direction of pitch axis.

[0033]

In addition, the pattern connecting unit 5 is

arranged to connect the unit patterns by interpolation with straight lines or with three dimensional curve according to classifications and the like of the used unit patterns.

[0034]

Next, operation of the pitch pattern generator with such configuration will be described. Firstly, a unit pattern is cut out from human utterance and this is stored in the pattern storing unit 1 in advance. Segmentation of unit patterns can be implemented with the methods as described above in Figure 2, Figure 3, Figure 4 or Figure 5 for example.

[0035]

That is, in case of using a method as in Figure 2, a unit pattern can be cut out as a unit the accent phrase from the starting point of the ascending portion to the ending portion of the descending portion via the plateau portion. As those similar hereto, direct section of the accent phrase as a unit may be considered, and also in this case, likewise Figure 2, the accent phrase pattern can be stored approximately in its entirety, but, compared with this, the way shown in Figure 2 is advantageous in the point that phrases with the same accent type but with different numbers of morae can be treated in the same way.

[0036]

In addition, in case of using such a method as in

Figure 3, unit patterns can be cut out of the accent phrase from the starting point of the ascending portion to the ending point of the successive plateau portion, and from the starting point of the descending portion to the ending point of the successive plateau portion respectively as a unit. As those similar hereto, section only of the ascending portion and the descending portion as a unit may be considered, and in this case as well, only by determining the location of the starting point of the unit pattern, the unit pattern can be disposed nearly completely, but as compared herewith, the way described in Figure 3 is advantageous in such a point that deterioration due to natural causes occurred by interpolation at the time of connection between patterns can be reduced.

[0037]

In addition, in the case where the way as in Figure 4 is used, the unit pattern can be cut out only with the plateau portion as a unit. The document "Study on fluctuation characteristics of basic cycle and on its model" by Osamu Komuro and Hideki Kasuya, The Journal of the Acoustical Society of Japan, Vol. 47, No. 12, pp. 928-934, 1991 reports that slow fluctuation in the constant vowel portion influences natural nature and section with the method in Figure 4 can preserve such fluctuation that is auditorily important.

[0038]

In addition, in the case where the way as in Figure 5 is used, a unit pattern can be cut out with such a characteristic pitch transforming portion as a unit that expresses the intonation of the end of a sentence. Figure 5 shows an example of pattern with an end rising intonation representing a question. According to the method in Figure 5, the pattern of the intonation at the end of a sentence can be stored in its entirety, even expression of difference in subtle nuance that is hard to express with rules will become feasible.

[0039]

The examples in Figure 2 to Figure 5, usage of the unit patterns cut out in relation with attribute information such as accent core location is premised, but otherwise, a method to use an automatically cut-out unit pattern with a pattern having percentage of change of pitch or curvature within a particular range may be considered. The former is advantageous in the point that the portion where the pitch shifts can be stored as a pattern and thus its outward appearance can be stored, and the constant domain is stored as a pattern so that fluctuation can be preserved. In addition, the latter stores a curve-like portion as a pattern and thereby the other portions can be interpolated with straight lines, process quantity can be reduced. As for the unit pattern cut out with methods in Figure 2

to Figure 5 or other methods, this may be directly stored in the pattern storing unit 1, or may be stored by coding subject to vector-quantization etc.

[0040]

Thus, after storing the unit pattern into the pattern storing unit 1, actual pitch pattern generating process can be started. In this pitch pattern generating process, firstly a predetermined input attribute is given to the pattern searching unit 2 which is made to search the unit pattern suitable to this input attribute from the pattern storing unit 1. Here, in the present embodiment, to the input attribute, beside the representative accent type and the like, at least one of information on classification of phonology in the vicinity before or after the pitch shifts, information of prominence, or intonation style of a phrase is added. In various items of information of such input attributes, in the case where a unit pattern is searched with phonological environment over a plurality of morae in the vicinity of the pitch shifting portion, the reproducing nature of local shifts of a pattern due to phonological environment is improved so that more precise pattern can be generated. In addition, in case of using prominence information to search unit patterns, intonations in synthesized voices which conventionally tend to become monotonous are made well-modulated so that intelligibleness and naturalness

can be improved. In addition in case of searching the unit patterns with intonation style, intonations which it is difficult to generate with rules can be accurately reproduced. In addition, in case of searching the unit patterns with neural networks, patterns close to speech patterns learned from a person can be reproduced so that naturalness can be improved.

[0041]

Thus, after an optimum unit pattern has been searched, the pattern disposing unit 3 determines a location of the searched unit pattern on the chronological axis in the generated pitch pattern based on the location mainly of accent core or accent phrase boundary on the chronological axis, and in more particular, determines the location of the starting point of the unit pattern to dispose the unit pattern. In more particular, the pattern disposing unit 3 determines the position of the starting point of the unit pattern in the phonological boundary location where pitch level on symbol level shifts in consideration of deviation which shifts due to phonological environments in that portion. Here, the location where the pitch level shifts is a phrase boundary or core location on accent phrases, and as for the unit pattern as shown in Figure 5, the location of a rising intonation representing a question, for example, is determined with the starting point of the

last vowel as a reference, which is known based on a document "Perception of Japanese questions", by Ichiro Miura and Miyoko Sugifuji, Acoustical Society of Japan autumn symposium lectures and papers, 1991, I, 2-6-9, pp255-256" and the like, and therefore, the boundary locations such as the phonology, syllables and morae etc. in the end of phrases are used. Figure 6 is a drawing to show how pattern disposing process looks like when the unit pattern as shown in Figure 3 is used, and in Figure 6, the bottom part represents the height of the pitch level in terms of the phonological boundary location and the signal level while the top part represents how the unit pattern is disposed. In Figure 6, the dotted line L1 represents the boundary between the ascending portion or the descending portion and the plateau portion, and in this base, the location of the starting point of the unit pattern is determined like the bold line L2.

[0042]

Thus, after the location of the unit pattern on the chronological axis, the pattern transforming unit 4 implements process of transforming the unit pattern so that it is embedded into the generated pitch pattern naturally. That is, at first the pattern transforming unit 4 proceeds with transformation on the chronological axis. Figure 7 is a drawing how the unit pattern subject to disposition as in Figure 6 undergoes

transformation. That is, in Figure 6, the unit pattern as shown in Figure 3 is used, and therefore in Figure 7, the starting point of each unit pattern is expanded and contracted in the direction of the chronological axis so as to correspond to the ending part of the preceding unit pattern. At this time, it is considered appropriate from the point of view of physiological requirements of vocal mechanism that expansion and contraction in the direction of chronological axis is implemented only in the plateau portion. Here, actual method of expansion and contraction will not be limited in particular, and expansion and contraction includes shutting down of, repetition of or extrapolation of data and the like.

[0043]

Thus, subject to expansion and contraction in the direction of chronological axis, transformation on the pitch axis is implemented. Figure 8 depicts a drawing how the pitch pattern subject to transformation in the direction of chronological axis undergoes transformation on the pitch axis. That is, this example undergoes expansion and contraction in the direction of pitch axis so that the starting point of each unit pattern corresponds with the ending point of the preceding unit pattern. At this time, it is appropriate that the ascending portion and the descending portion shift according to preceding and

succeeding context, expansion and contraction in the direction of pitch axis is implemented well only in the ascending portion as well as the descending portion, and thereby, the rough shape or fluctuation in the plateau portion can be preserved. Such a method preserves the plateau portion of the unit pattern also in terms of its absolute height. Thus, including influence of prominence or influence due to location in a phrase, the pitch pattern can be reproduced with a unit pattern.

[0044]

In the above described example, the unit patterns as shown in Figure 3 are used, and therefore on pattern-to-pattern connection interpolating process is not needed in particular, but in case of using the unit patterns other than the unit patterns as shown in Figure 3, certain interpolating process will be required for the pattern connecting unit 5. At this time, in case of adopting only the ascending portion and the descending portion only as unit patterns, the gap in-between is interpolated with a straight line and in case of adopting the plateau portion as the unit pattern as shown in Figure 4, the gap in-between is interpolated with a three dimensional curve so that patterns closer to human pitch pattern can be generated.

[0045]

Here, in a pitch pattern generator of unit

connection type as described above, a problem which unit pattern should be stored in the pattern storing unit significantly influences its performance. Therefore, further in the present invention, as the unit pattern stored in the pattern storing unit, pitch patterns characteristically appearing under a particular linguistic situation is used, and thereby, efficient and effective pitch pattern generation is arranged to be realized. Such unit patterns will be described in detail as follows.

[0046]

Continuous utterance includes a plurality of phrases such as breath group and sentence, and as the location in a phrase is getting close to the end, normal speech will take low pitch patterns. Accordingly, preparing a plurality of different patterns in in-phrase locations, without providing complicated process such as separation/synthesis in a beginning component, will give rise to the phrasal mood described above. In addition, regardless the sentence length at the end of a sentence, use of comparatively low specific-purpose pattern can give rise to a statement-like mood. Thus, as one of the unit patterns, use of pitch patterns characteristically appearing in a particular location in a series of phrases such as breath group and sentences can express a phrasal mood and a statement-like mood.

[0047]

In addition, rough shapes of patterns of accent phrases are different according to existence or locations of accent cores. Coreless accent phrase pattern is smoother than that of a phrase with a core, and a initial-high type ascends a little steeply and descends gradually. Accordingly, selective use of a plurality of patterns with different accent types can express such differences. Thus, as one of unit patterns, use of pitch patterns characteristically appearing in a particular accent type such as an initial-high type and a plateau type can generate more natural pitch patterns.

[0048]

In addition, using a comparatively high special-purpose pattern, the portion provided with prominence can give rise to natural prominence mood without any unnatural pitch ascend even in the head of a phrase. In addition, expression of subsidiary accent which was conventionally difficult will become feasible by preparing patterns with different accent levels. Thus, as one of unit patterns, using pitch patterns which characteristically appear in a particular accent levels such as in a portion provided with prominence and in a portion where a subsidiary accent core appears, prominence and subsidiary accent will become expressible and naturalness can be improved.

[0049]

In addition, as one of the unit patterns, words which are uttered with their specific pitch patterns such as in interjection, an end particle, an onomatopoeic word, and imitative word and a shout are used, or phrasal pitch patterns can be directly used, and in this case, words which are uttered with their own characteristic pitch patterns can be reproduced faithfully, and naturalness can be improved.

[0050]

In addition, in case of reading numbers plainly or in case of reciting a verse such as Waka, a constant pitch pattern appears periodically. Accordingly, as one of the unit patterns, use of characteristic pitch patterns which periodically appear in a line of numbers and a verse can cope with utterance with their own intonations as described above.

[0051]

In addition, in conversation, as in "Eh" and "Ng," etc., the words themselves do not carry meaning, intentions are frequently expressed mainly with difference in intonations. Accordingly, as one of unit patterns, use of characteristic pitch patterns expressing particular intentions mainly by way of meter independent from words and phonology enables efficiently natural conversation output.

[0052]

Figure 9 is a diagram showing a specified example on process to search these unit patterns from input linguistic information to synthesize pitch patterns. In the example in Figure 9, the pattern storing unit 1 shows the case where seven kinds of unit patterns of [P01] initial-plateau type, [P03] end plateau type, [P41] prominence, [P42] subsidiary accent, [P51] is it?, [P61] number, [P71] unexpected are stored in advance, and when "e? (unexpected)/is it 920917? once again/menu number (prominence)/please input" is inputted as input linguistic information, the pattern searching unit 2 searches a unit pattern column corresponding with this input linguistic information from the pattern storing unit 1. As a result, the unit pattern column of "[P71] e? [P61]92 [P61]09 [P61]17 [P51]is it? [P01]once again [P41] menu number [P03]input [P42]please" can be searched. Next, thus searched unit pattern column is brought into disposition, transformation and connection in the pattern disposing unit 3, the pattern transforming unit 4 and the pattern connecting unit and thereby a natural pitch pattern as shown in the bottom of Figure 9 can be generated.

[0053]

Thus, as the unit pattern stored in the pattern storing unit 1, pitch patterns characteristically appearing under a particular linguistic situation is used, and from input linguistic information, the unit

patterns can be searched heuristically so that these unit patterns are brought into disposition, transformation and connection to make generation of natural pitch patterns possible efficiently and effectively with comparatively simple process.

[0054]

Here, in case of directly storing pitch patterns of natural sounds as the unit patterns, quite a few data larger than those to be stored as generation rules or model parameters will become necessary and the real-time process is expected to become difficult due to increase in memory volume as well as increase in process volume. Therefore, something to avoid it should be devised.

[0055]

In order to prevent memory volume as well as process volume from increasing and thereby making real-time process difficult, the inventor of the present application has further invented to approximate the unit patterns with line to store in the pattern storing unit 1 without storing the unit patterns cut out in an appropriate unit from pitch patterns of natural voices in the pattern storing unit 1 in their original forms.

[0056]

Figure 10 shows an example of having approximated by line the unit patterns which were originally curved. In the example of Figure 10, one unit pattern is

approximated by seven lines, and are divided into section (segment) S1 to S7 for each line. Here, a unit pattern can be expressed with a patch at the staring point, section length (segment length) of respective section, and a step value expressing inclination of line (a line portion) of respective section, and the unit pattern can be stored in this expression form in the pattern storing unit 1. Here, a step value refers to a time period required for pitch cycle for one sampling transformation, that is, a pitch renewal cycle, providing positive (+) values for the line inclined to ascend rightward as in Figure 11(a) and providing negative (-) values for the line inclined to descend rightward as in Figure 11(b).

[0057]

Thus, approximating the originally curved unit patterns with a line, which is expressed with the pitch at the starting point as well as the segment length in the respective section and the step value to be stored in the pattern storing unit 1, the data quantity stored in the pattern storing unit 1 can be dramatically reduced, compared with the case of storing the unit patterns in their original forms. Moreover, shift in the pitch direction of a unit pattern can be implemented only with increase/decrease in the pitch at the starting point, and expansion and contraction in the time axis direction can be implemented with

increase/decrease in segment length and transformation of pattern can be implemented simply.

[0058]

However, in the case where the expression as described above is used, when the pattern searching unit 2 has searched and read out the data in the above described expression form of a predetermined unit pattern from the pattern storing unit 1, it is necessary to reproduce (generate) this into the unit pattern form (that is, into line shape), and the unit pattern generating process can be realized with a simple configuration as shown in Figure 12. That is, a circuit shown in Figure 12, to which for reproducing (generating) a line in a certain section the step value in that section will be inputted, is configured by comprising an inverter 11 for obtaining the absolute value of the inputted step value (that is, the negative sign is reversed when the step value is negative), a step counter 12 for outputting "1" when the absolute value of the step value has been set, the sample clock in a predetermined time interval is counted, and the counted value of the sample clock has corresponded to the absolute value of the step value, a comparator 13 for determining whether the inputted step value is positive or negative, an AND circuit 14 for outputting "1" when the comparator 13 has determined as positive (+) and "1" has been outputted from the step counter 12,

an AND circuit 15 for outputting "1" when the comparator 13 has determined as negative (-) and "1" has been outputted from the step counter 12, a pitch counter 16 for counting the output "1" from the AND circuit 14 in the positive (+) direction, counting the output "1" from the AND circuit 15 in the negative (-) direction and outputting that count value as a pitch, and with this circuit, process of increasing the pitch cycle by unit length in case of the step value (pitch renewing cycle) being positive and of decreasing the pitch cycle by unit length in case of the step value being negative can be implemented. In other words, setting the pitch at the starting point as the initial value of the pitch counter 16, causing the pitch counter 16 in the above described circuit to keep outputting pitches until a section length (segment length) in the time axis direction in one section, this is implemented sequentially on each segment, and thereby unit patterns in approximation by line can be reproduced (generated). Thus, unit patterns can be reproduced, and actual pitch patterns can be generated.

[0059]

In addition, the pattern disposing unit 3 disposes at a predetermined location each unit pattern searched and generated by the pattern searching unit 2, and in order to obtain natural pitch patterns, the pattern transforming unit 4 implements process to transform

each unit pattern disposed by the pattern disposing unit 3. At this time, due to connection between adjacent unit patterns and provision of prominence, it is necessary to adjust inclination. In the present invention, since data expression as described above is adopted, this adjustment treatment can be easily realized by, for example, adding a constant bias to the step values over all the sections of the transitional portion with an (not shown) adding machine (to refer to Figure 10).

[0060]

Figure 13 shows an example of such an adjusting process for one unit pattern. With reference to Figure 13, adding a constant bias value respectively to the step value of each section of one unit pattern, and changing the step value (changing inclination of each line), smooth transformation in the pitch direction can be provided to the unit patterns. For example, as in the example of Figure 13, adding negative bias value BIAS to the step value (prior to transformation) ST_A of each section of the unit pattern to get ST_B, the unit pattern UNP_A prior to transformation can be transformed as UNP_B. That is, without giving rise to any change as for the ending point, a predetermined quantity (of bias quantity) of change can be given rise to the starting point. Thus, a pattern transforming process that will require connection between the unit patterns and

provision of prominence can be realized with a very simple configuration only with an adding machine.

[0061]

In order to adjust natural inclination in such a way, it is necessary to cause the right ascending inclination to horizontal or right descending inclination successively, and this will be feasible by causing the maximum value (positive maximum absolute value) and the minimum value (negative maximum absolute value) of the step value to change successively, and such process can be easily realized by treating the step value (pitch renewal cycle) with complement expression of "2" to, in the case where the absolute value of the step value as a result of addition of the bias surpasses the maximum value, regard this as the maximum absolute value with the opposite signal.

[0062]

Thus, the pattern transforming unit 4 applies a constant bias to the unit value of each section of one unit pattern so that the unit pattern can undergo smooth transformation in the pitch direction, but the pattern transforming unit 4 must calculate on a real-time basis the bias which will become indispensable for connection between adjacent unit patterns or for provision of prominence, etc. This calculation normally requires multiplying and dividing calculation to make the process complicated, which could be the

bottleneck against real-time implementation of the pitch pattern generating process in its entirety. In order to avoid this problem, calculating in advance the corresponding relationship between the pitch shifting quantity (for example, the balance between the pitches at the starting point respectively of the original pattern and the generated pattern, or the pitch itself at the starting point of the generated pattern) and the bias to provide each unit pattern with a bias table as shown in Figure 14, and when a certain unit pattern undergoes transformation at a predetermined pitch shifting quantity in the pitch direction, a bias corresponding to the above described pitch shifting quantity is successfully read out for use from this unit pattern bias table. This deprives necessity of calculating bias for each time, without giving rise to troubles to real-time process. In particular, as shown in Figure 15, as for the ending point of each pattern, this is not caused to shift, but calculating the pitch shifting quantity at the starting point so that the starting point of each pattern will become the ending point of the preceding pattern, the bias corresponding to this is read out for use from the bias table as shown in Figure 14 to transform the original pattern to constitute a generated pattern, and thereby connection with the preceding pattern can be implemented easily.

[0063]

In addition, as a variation of the present invention, in order to making maintenance such as addition, deletion, change and the like of the unit pattern in the pattern searching unit 2 easy, as shown in Figure 16, a rule storing unit 7 which stores a plurality of rules for searching unit patterns on an input linguistic information unit base can be provided further.

[0064]

In this case, as rules stored in the rule storing unit 7, the one configured by conditions consisting of a plurality of items and a range of values thereof, and unit numbers can be used. Here, the items are respective kinds of attributes on each mora, and in particular, locations from the phrase top, boundary categories with adjacent phrases, accent types, accent levels, phonology and the like are deemed effective.

[0065]

Next, particular example of pattern searching treatment in the pitch pattern generator configured as shown in Figure 16 will be described with reference to Figure 17 to Figure 19. Figure 17 depicts a table showing respective kinds of unit patterns stored in the pattern storing unit 1, and in the example in Figure 17, four kinds of unit patterns are provided with unit numbers "1", "2", "3" and "4" and stored. In addition, Figure 18 depicts a table showing a configuration

example of rules stored in the rule storing unit 7, and the example in Figure 18 adopts as items four attributes of the location POS from the accent phrase top, the location Mora from the beginning phrase top, an accent type Acc, the location Core of the accent core, and value ranges of four attributes are allotted respectively to four kinds of unit numbers "1", "2", "3" and "4".

[0066]

In Figure 16, the pattern searching unit 2 refers to the rule storing unit 7 in Figure 18 to search the unit patterns corresponding with the input linguistic information from four kinds of unit patterns stored in the pattern storing unit 1. In particular, taking the mora as input linguistic information, and considering the process to take place on a mora unit basis, in the case where input mora column as shown in Figure 19 is inputted, at first the values for all items are obtained on each mora. Focusing on a certain mora, the pattern searching unit 2 refers to the rule storing unit 7, looks into all items included in the rule conditions to confirm whether or not the values of each item of this mora are included in that range. And, when the value of each item of this mora is judged to exist in the range of items of a certain unit number, this mora will be regarded to fulfill the rule conditions, and the pattern searching unit 2 can search

from the pattern storing unit 1 a unit pattern corresponding to the unit number of that rule. In addition, when there is no rule fulfilling the conditions, no unit pattern to be disposed in that mora location exist, and the process goes to the next mora.

[0067]

For example, in the input mora column in Figure 19, the mora "ro" in "meron" comprises respective item values of "0", "4", "1" and "0", and fulfills the rule conditions of the unit number "4" among the rules stored in the rule storing unit 7, and therefore the pattern searching unit 4 can search the unit pattern of the unit number "4" from the pattern storing unit 1.

[0068]

Thus, causing the rule conditions to include a range of values of a plurality of variables representing characteristics of linguistic information unit, the rule searching can be implemented with routine process. In addition, it will become easy to separate the rules as external data so that maintainability of the rules themselves can be improved.

[0069]

In addition, usage of locations from the phrase top or the phrase end of the linguistic information unit as the variables included in the rule conditions will enable expression of nature of pitch patterns continuously changing in accordance with locations in

the phrase such as gradual descend from the phrase top to the phrase end. In addition, usage of classification of boundary with the phrase adjacent to the phrase containing the linguistic information unit will enable expression of pitch patterns in the phrase boundary locations such as direction on ending impression of a statement by dropping the statement end in particular lower than the period position, and dealing with various intonation styles such as questioning statement. In addition, usage of accent information of the phrase containing the linguistic information unit will enable expression of pitch pattern characteristics such as characteristics according to accent type such as the initial rise type constituting a mountain leftward-inclined than the middle rise type, etc. and characteristics according to prominence or subsidiary accent and the like. In addition, usage of phonological information of the linguistic information unit will enable expression of pitch pattern characteristics locally observed in a particular phonology or influence to phonology sustaining chronological length and the like.

[0070]

In the above described example, the rule and the unit pattern is to constitute one-to-one correspondence, but one rule could have a plurality of unit numbers. In this case, some "fluctuation" is provided so that

one unit pattern may be selected. Thereby, the same input can generate a totally dissimilar pitch pattern so that uniform and mechanical pitch patterns will become avoidable and more natural pitch patterns can be obtained.

[0071]

[Effects of the Invention]

As described so far, according to the invention according to claim 1, unit patterns cut out on an appropriate unit basis from pitch patterns of human speeches are stored in pattern storing means, pattern searching means searches unit patterns for conforming to the input attribute based on a predetermined input attribute, pattern disposing means disposes the unit patterns obtained by searching onto the chronological axis, and pattern transforming means transforms the unit patterns disposed on the chronological axis so that pitch patterns are arranged to be generated, and since the pitch patterns that a human being has spoken are directly used, such pitch patterns that will be able to express natural variations, fluctuation, nuance and the like that are difficult for rule-based generation can be generated.

[0072]

In addition, in the invention according to claim 2, in case of cutting out the one containing one ascending portion and one descending portion of accent phrase

respectively as a unit that is to be used as the unit pattern, pitch patterns with the accent phrase patterns in nearly full preservation being important units of pitch patterns can be generated. In addition, in case of using the one containing one ascending portion or one descending portion of accent phrase as the unit pattern, pitch patterns can be generated with a simple connecting process. In addition, in case of using the unit patterns containing plateau portions of accent phrase, fluctuation and the like in the constant domain can be preserved and natural pitch patterns can be generated. In addition, in case of using characteristic pitch variations in the end of phrases as the unit patterns, various intonations which it is difficult to generate with rules can be expressed. In addition, in case of using patterns showing percentage of change of pitch within a particular range as the unit patterns, the rough form in shifting portions and fluctuation etc. in the constant domain portions can be preserved so that various and natural pitch patterns which it is difficult to generate with rules can be generated. In addition, in case of using patterns showing pitch curvatures within a particular range as the unit patterns, since the other portions can be interpolated with straight lines, process quantity can be reduced.

[0073]

In addition, in the invention according to claim 3, the unit patterns subject to vector quantization are stored in the pattern storing unit, and when this is used, the volume of pattern storing means can be dramatically reduced so that cost reduction as well as miniaturization in the device and increasing in processing speed can be achieved.

[0074]

In addition, according to the invention according to claim 4, searching the unit patterns based on phonological environments over a plurality of morae in the vicinity of pitch shifting portions, the reproducing nature of local shifts of patterns due to phonological environments is improved so that more precise pitch patterns can be generated. In addition, in case of searching the unit patterns based on prominence information, intonations in synthesized voices which conventionally tend to become monotonous are made well-modulated so that intelligibleness and naturalness can be improved. In addition, in case of searching the unit patterns based on intonation styles, intonations which it is difficult to generate with rules can be accurately reproduced.

[0075]

In addition, according to the invention according to claim 5, searching the unit patterns with neural networks, patterns close to speech patterns learned

from a person can be reproduced so that naturalness can be improved.

[0076]

In addition, in the invention according to claim 6, in case of disposing the unit patterns with locations of accent phrase boundary, the ascend starting points of accent phrases can be reproduced to an approximately full extent so that intelligibleness can be improved. In addition, disposing unit patterns with locations of accent cores, the descend starting points of accent phrases can be reproduced to an approximately full extent so that intelligibleness can be improved. In addition, disposing unit patterns with boundary location in the end of phrases such as phonology, syllables and morae, etc., intonations can be expressed efficiently so that naturalness can be improved. In addition, disposing the unit patterns with phonological environment information in the vicinity of pitch shifting portions, chronological structure of pitch patterns can be reproduced more precisely so that naturalness can be improved.

[0077]

In addition, according to the invention according to claim 7, the plateau portions in the accent phrases are to undergo expansion and contraction in the direction of chronological axis so that, as for the pitch shifting portions, without any shift in the

direction of chronological axis, the chronological structure of pitch patterns can be preserved and the patterns which are not against physiological restrictions are guaranteed and therefore stable natural pitch patterns can be generated. In addition, the ascending portions and the descending portions in the accent phrases are to shift in the direction of pitch axis so that, as for the plateau portions, without any shift in the direction of the pitch axis, height or fluctuation etc. of pitch constant domain can be preserved, pattern expressions including prominence and the like become feasible and naturalness can be improved.

[0078]

In addition, according to the invention according to claims 8 and 9, interpolating interconnection between the unit patterns with straight lines, process quantities can be reduced and increase in processing speed as well as cost reduction can be achieved while, in case of interpolating interconnection between the unit patterns with three dimensional curve, rough shape similar to human pitch shifts can be expressed and natural pitch patterns can be generated.

[0079]

In addition, according to the invention according to claim 10, since the pitch patterns characteristically appearing in a particular linguistic

situations are arranged to be used as the unit patterns stored in the pattern storing unit, patterns can be searched heuristically from input linguistic information so that efficient and effective pitch pattern can be generated.

[0080]

In addition, according to the invention according to claim 11, as one of the above described unit patterns, use of the pitch patterns characteristically appearing in a particular location in a series of phrases such as breath group and sentences enables expression of a phrasal mood and a statement-like mood without providing complicated process such as separation/synthesis in a beginning component so that simplification of process can be achieved.

[0081]

In addition, according to the invention according to claim 12, as one of the above described unit patterns, using the pitch patterns characteristically appearing in a particular accent type such as a initial-high type and a plateau type, rough shapes of patterns of accent phrases which are different according to existence or locations of accent cores can be expressed by selective use of a plurality of patterns with different accent types and naturalness can be improved.

[0082]

In addition, according to the invention according to claim 13, as one of the above described unit patterns, using the pitch patterns which characteristically appear in a particular accent levels such as in a portion provided with prominence and in a portion where a subsidiary accent core appears, prominence and subsidiary accent will become expressible and naturalness can be improved.

[0083]

In addition, according to the invention according to claim 14, as one of the above described unit patterns, words which are uttered with their specific pitch patterns such as in interjection, an end particle, an onomatopoeic word, and imitative word and a shout are used, or phrasal pitch patterns are directly used, and thereby, words which are uttered with their own characteristic pitch patterns can be reproduced faithfully, and naturalness can be improved.

[0084]

In addition, according to the invention according to claim 15, as one of the above described unit patterns, use of characteristic pitch patterns which periodically appear in a line of numbers and a verse can cope with utterance with their own characteristic intonations such as in case of reading numbers plainly or in case of reciting a verse, improving general-purpose application of the device.

[0085]

In addition, according to the invention according to claim 16, as one of the above described unit patterns, use of the characteristic pitch patterns expressing particular intentions mainly by way of meter independent from words and phonology enables efficiently natural conversation output also in case of expressing intentions mainly with difference in intonations such as "Eh" and "Ng," etc. which are characteristic in conversations.

[0086]

In addition, the inventions according to claims 17 and 18 are to arrange the pattern storing means to store the unit patterns subject to approximation with line and, at this time, the unit patterns subject to approximation with line are expressed with expression data with the pitch at the starting point and the inclination between respective section of each line and the line, and stored in the pattern storing means, thereby data quantities of the unit patterns memory amount and process amount can be reduced and cost reduction as well as increase in processing speed can be achieved. Moreover, shifting in the pitch direction of the unit patterns is only increase or reduction of staring pitch, and expansion and contraction in the chronological direction can be simply realized respectively with increase or reduction of interval

length, thereby unit pattern shifting can be realized with simple process so that increase in processing speed as well as functional improvement can be achieved.

[0087]

In addition, according to the invention according to claim 19, inclination of each of the above described line is expressed by step values as time required for the pitch cycle to change by a unit quantity, and in this case, when the pattern searching means searches expression data of the unit patterns stored in the pattern storing means and reproduces the unit patterns based on the expression data the pattern searching means implements process to increase the pitch cycle by the unit period in case of the step values being positive values and decrease the pitch cycle by the unit period in case of the step values being negative values, and therefore extremely simple configurations with counters and comparators etc. can reproduced the unit patterns approximated with lines so that actual pitch patterns become obtainable.

[0088]

In addition, according to the invention according to claim 20, pattern transforming means adds a constant bias to the step value representing inclination of each line so as to provide the unit patterns with smooth shifting in the pitch direction, and therefore the pattern transforming process that will become necessary

for connection between the unit patterns or provision of prominence can be realized with extremely simple configurations only with an adding machine.

[0089]

In addition, according to the invention according to claim 21, the pattern transforming means treats the step values with complement expression of "2" and, in the case where the absolute value of the step value as a result of application of bias has transcended the maximum value, regard this as the maximum absolute value with opposite sign, and therefore a process to implement pattern transforming continuously and without compulsiveness from right ascending pattern to right descending pattern or vice versa can be realized simply.

[0090]

In addition, according to the invention according to claim 22, in the case where corresponding relationship between the pitch transforming quantity and the bias is obtained in advance, the pattern transforming means implements shifting in the pitch direction of the unit pattern with the corresponding relationship, and therefore the bias that will become necessary in connection between unit patterns as well as prominence application and the like can be obtained extremely easily without necessity of multiplication and division with a large amount of process quantity so that increase in processing speed can be achieved.

[0091]

In addition, according to the invention according to claim 23, moreover, the rule storing means for storing rules including conditions corresponding with the unit patterns stored in the pattern storing means is provided so that the pattern searching means is arranged to look into the rule conditions stored in the rule storing means in each input linguistic information unit such as morae etc. and thereby searches the unit patterns fulfilling the rule conditions from the pattern storing means, and adopting the rule searching as the unit pattern searching process, maintenance such as addition, deletion and change etc. in unit patterns will become easy.

[0092]

In addition, according to the invention according to claim 24, the rule conditions include ranges of values of a plurality of variables representing characteristics of linguistic information unit, and the pattern searching means is arranged to adopt the rule when for all the variables included in the rule conditions the values of the input linguistic information unit are within that range, and thereby to search the unit patterns, and making the rule conditions include a range of values of a plurality of variables representing characteristics of linguistic information unit, the rule searching can be implemented

with routine process. In addition, it will become easy to separate the rules as external data so that maintainability of the rules themselves can be improved.

[0093]

In addition, according to the invention according to claim 25, as the variables included in the rule conditions, locations from the phrase top or the phrase end of the linguistic information unit are used, classification of boundary with the phrase adjacent to the phrase containing the linguistic information unit is used, accent information of the phrase containing the linguistic information unit is used or phonological information of the linguistic information unit is used, and usage of locations from the phrase top or the phrase end of the linguistic information unit as the variables included in the rule conditions will enable expression of nature of pitch patterns continuously changing in accordance with locations in the phrase such as gradual descend from the phrase top to the phrase end. In addition, usage of classification of boundary with the phrase adjacent to the phrase containing the linguistic information unit will enable expression of pitch pattern characteristic in the phrase boundary locations such as direction on ending impression of a statement by dropping the statement end in particular lower than the period position, and dealing with various intonation styles such as

questioning statement. In addition, usage of accent information of the phrase containing the linguistic information unit will enable expression of pitch pattern characteristics such as characteristics according to accent type such as the initial rise type constituting a mountain leftward-inclined than the middle rise type, etc. and characteristics according to prominence or subsidiary accent and the like. In addition, usage of phonological information of the linguistic information unit will enable expression of pitch pattern characteristics locally observed in a particular phonology or influence to phonology sustaining chronological length and the like.

[0094]

In addition, according to the invention according to claim 26, a plurality of unit patterns are brought into correspondence with one rule, and, in the case where a plurality of unit patterns fulfilling the rule condition are searched, fluctuation is provided so that one unit pattern will be selected, and therefore uniform and mechanical pitch patterns will become avoidable and more natural pitch patterns can be obtained.

[Brief Description of the Drawings]

[Figure 1] Figure 1 is a block diagram of an embodiment of a pitch pattern generator according to the present invention.

[Figure 2] Figure 2 is a diagram for describing an example of cutting out unit patterns.

[Figure 3] Figure 3 is a diagram for describing an example of cutting out unit patterns.

[Figure 4] Figure 4 is a diagram for describing an example of cutting out unit patterns.

[Figure 5] Figure 5 is a diagram for describing an example of cutting out a unit pattern.

[Figure 6] Figure 6 is a diagram for describing how locations of unit patterns on the chronological axis are determined.

[Figure 7] Figure 7 is a diagram for describing transformations of unit patterns on the chronological axis.

[Figure 8] Figure 8 is a diagram for describing transformations of unit patterns on the pitch axis.

[Figure 9] Figure 9 is a diagram for describing treatment to search unit patterns from input linguistic information to synthesize the pitch patterns.

[Figure 10] Figure 10 is a diagram for describing an example where unit patterns have undergone approximation by line.

[Figure 11] Figure 11 (a) and (b) are diagrams for describing step values.

[Figure 12] Figure 12 is a diagram for describing an example of pitch pattern generator.

[Figure 13] Figure 13 is a diagram for describing

process to adjust inclinations of unit patterns.

[Figure 14] Figure 14 is a table showing an example of bias table.

[Figure 15] Figure 15 is a diagram for describing pattern transformation process.

[Figure 16] Figure 16 is a diagram for describing an example of variation of pitch pattern generator shown in Figure 1.

[Figure 17] Figure 17 is a diagram for describing respective kinds of unit patterns stored in the pattern storing unit.

[Figure 18] Figure 18 is a table showing an example of configuration of rules stored in the rule storing unit.

[Figure 19] Figure 19 is a drawing showing an example of input mora columns.

[Description of the Symbols]

1 pattern storing unit

2 pattern searching unit

3 pattern disposing unit

4 pattern transforming unit

5 pattern connecting unit

7 rule storing unit

Figure 1

1 pattern storing unit

2 pattern searching unit

3 pattern disposing unit

4 pattern transforming unit

5 pattern connecting unit

#1 input attribute

#2 pitch pattern

Figure 6

#1 pitch axis direction

#2 chronological axis

#3 phonological boundary

#4 symbolic accent level

#5 me

ro

n

ga

o

i

shi

i

Figure 7

#1 pitch axis direction

#2 chronological axis

#3 phonological boundary

#4 symbolic accent level

#5 me

ro

n

ga

o

i

shi

i

Figure 8

#1 pitch axis direction

#2 chronological axis

#3 phonological boundary

#4 symbolic accent level

#5 me

ro

n

ga

o

i

shi

i

Figure 9

1 pattern storing unit

[P01] initial-plateau type

[P03] end plateau type

[P41] prominence

[P42] subsidiary accent

[P51] is it?

[P61] number

[P71] unexpected

2 pattern searching unit

3 pattern disposing unit

4 pattern transforming unit

5 pattern connecting unit

#1 input linguistic information

#2 e? (unexpected)/is it 920917?

once again/menu number (prominence)/please input

#3 searched unit pattern column

#4 [71] e? [P61]92[P61]09[P61]17[P51] is it?

[P01] once again [P41] menu number [P03] input

[P42] please

#5 generated pitch pattern

#6 [P71] e?

#7 [P51] is it?

#8 [P01] once again

#9 [P41] menu number

#10 [P03] input

#11 [P42] please

Figure 10

#1 pitch at starting point

#2 segment length

#3 step value

#4 transitional portion

#5 constant domain portion

Figure 11

#1 step value

Figure 12

#1 sample clock

#2 step value

#3 pitch

12 step counter

13 comparator

16 pitch counter

Figure 14

#1 pitch shifting quantity

#2 bias

Figure 15

#1 preceding pattern

#2 generated pattern

#3 original pattern

#4 bias

Figure 16

1 pattern storing unit

2 pattern searching unit

3 pattern disposing unit

4 pattern transforming unit

5 pattern connecting unit

7 rule storing unit

#1 input linguistic information

#2 pitch pattern

Figure 17

#1 unit number

#2 unit pattern

Figure 18

#1 unit number

Figure 19

#1 a

ma

i

me

ro

n

o

ta

be

ma

shi

ta